
THE PROPOSITION VALUE OF CORPORATE RATINGS - A RELIABILITY TESTING OF CORPORATE RATINGS BY APPLYING ROC AND CAP TECHNIQUES

LIS Bettina

University of Mainz, Germany

NEßLER Christian

University of Mainz, Germany

RETZMANN Jan

Deutsche Pfandbriefbank, Germany

Abstract:

We analyze the Altman model, a Logit model as well as the KMV model in order to evaluate their performance. Therefore, we use a random sample of 132 US firms. We create a yearly and a quarterly sample set to construct a portfolio of defaulting and a counter portfolio of non-defaulting companies. As we stay close to the recommendations of the Basel Capital Accord framework in order to evaluate the models, we use Receiver Operating Characteristic (ROC) and Cumulative Accuracy Profile (CAP) techniques. We find that the Logit model outperforms the Altman as well as the KMV model. Furthermore, we find that the Altman model outperforms the KMV model, which is nearly as accurate as a random model.

Keywords: Altman Model, Cumulative Accuracy Profile (CAP), Distance to Default, Logit Model, Moody's KMV, Receiver Operating Characteristic (ROC), Z-score.

1. Introduction

'For the second time in seven years, the bursting of a major asset - bubble has inflicted great damage on world financial markets.' While reading economical newspapers, one could find phrases like the one from Stephen S. Roach (Morgan Stanley) in nearly every kind of newspaper. The past crisis found its starting point with defaulting US consumer credits and thus affected banks capital requirements immediately.

The stock market reacted with massive price fluctuations. Especially bank and insurance titles got under enormous pressure. As a reaction, the European Central Bank (ECB) gave several short time credits to banks to secure liquidity. According to the Manager Magazine (2008), these credits amount to €94.841 billion at the 09.08.2007, €42.245 billion at the 06.09.2007 and €300 billion at the 25.09.2008. The FAZ (2008) reported that the credit crisis causes a deceleration in economical growth especially in the US but in the EU as well. Was it possible to forecast the crisis? The President of the ECB Jean-Claude Trichet agreed to this opinion and blamed rating agencies to embellish the situation by giving overvalued rating grades to high risk financial products (FTD (2007)). To evaluate companies and financial products, rating agencies are using different kinds of rating models. Typically these models evaluate default risk by categorizing the company / the financial product in a predefined rating scale. In general, a rating grade is a synonym for a default probability forecasting a time horizon of one year. However, the procedure of how these models work is mostly unknown. In addition to commercial rating models, academic literature offers a huge range of publicly available rating models. The Z-score model by Altman (1968) for example, is probably the most known rating model. This model heralds an era of new valuation models, using statistics in order to measure and describe a company's probability of default. Up to date the model is used as a benchmark for every kind of credit risk model. To compensate disadvantages of Altman's linear model, academic literature describes a huge range of models using other, non linear techniques. Staying close to the present discussion about the performance of rating models, we analyze models which are applicable within the Basel framework. (Basel Committee on Banking Supervision (2001)). Therefore, the aim and objective of this paper is to figure out, whether the Z-score model, the bounded Logit model as well as the KMV model are appropriate systems to measure a company's default risk. Dealing with the various models performances, Engelmann et al. (2003) describe, that a rating system's quality results from its discriminate power to correctly distinguishing between non-defaulting firms and defaulting firms forward looking for a predefined time horizon. In order to test the rating models correctness, we apply the 'Cumulative Accuracy Profile Model' (CAP) and the 'Receiver Operating Characteristics Model' (ROC) techniques. According to Engelmann, both techniques are the most accepted evaluation techniques currently used in practice in order to analyze rating models performance.

Due to the current developments and discussions regarding the regulation of rating agencies and the correspondent liability the analysis gets certain relevance (Eisen 2008). Applying these techniques, we find that the models differ in their forecast quality. Therefore, the Logit model outperforms the Altman as well as the KMV model. Furthermore, we find that the Altman model outperforms the KMV model. The results give helpful suggestions for understanding the models and the arrangement of the regulative institutions.

2. Model review

The literature and academic examination regarding credit risk modeling has increased immensely since the source works of Altman (1968) and Merton (1974). Due to liberalization of capital markets, increases liquidity, completion on bond markets, developing and revisions of Basel Capital Accord Framework there is a grown practitioner's attentiveness in correct credit risk assessment (Fernandes 2005). Structural models can be used for companies with equity or debt which is traded on markets. In that case the approaches of Black/Scholes (1973) or Merton (1974) or the extension by Black/Cox (1976) or Longstaff/Schwarz (1993) can be used. By using valuation methods of option pricing theory, a credit facility is seen as a contingent claim on the value of the company's assets. The default is defined to happen when the company hits the pre-defined default barrier. The Intensity models developed for e.g. by Jarrow/Turnbull (1995) or Duffie/Singleton (1997) do not try to measure the company's market value. The default is set to occur as the time of a first jump of a poisson process with random intensity. For companies which are not traded and no market based data is available, accounting related credit scoring models seem to be the common approach and following Allen (2002) seems to be the most effective and broadly accepted conceptualization (Fernandes 2005). The credit scoring approach was discussed intensively in academic literature since Beaver (1966) and Altman (1968). Barniv/McDonald (1999) state in a meta-analytic study that in the Logit models were used or discussed 178 times between 1989 and 1996 and underline the popularity of the Logit Model.

Saunders et al. (2002) describe that the Altman and the Logit model belong to the group of scoring systems. All scoring systems have in common that they use pre-identified weighted factors to determine the probability of default. Altman (2002) refers to the Z-score model and the Logit model by writing that both models have in common, that they involve a set of financial indicators in combination with qualitative elements. Financial ratios therefore analyze firm's profitability, liquidity as well as its solvency in order to forecast its wealth. Ratios as well as weights result from empirical observations so that they best distinguish between bankrupt and non-bankrupt firms for an underlying dataset. The models differ in the way how they process these ratios. Whereas the Z-score model processes them in a linear way, the Logit model processes them in a Logit function. The KMV approach differs substantially as it is based on option theory. According to Navneet et al. (2005) the bankruptcy takes place, when the firm's market value of debt exceeds the firm's market value of equity. In that case the entrepreneur strikes and transfers the firm's assets to the bank. In the following the models in the analysis (Z-score Model, the Bounded Logit Model and the KMV Model) get a brief definition.

2.1. The Z-score Model

Altman's (1968) Z-score model forecasts corporate bankruptcy based on weighted financial ratios, processed in a linear function. Altman criticizes the inaccuracy of pure ratios analysis in order to evaluate companies' default risk. He argues that especially size effects would deform the accuracy of ratios. The size effect explains that financial ratios deflate statistics by size. According to Altman, this is a particular problem if ratios are getting compared among different companies. In order to deal with the impact of size, Altman concentrates on multiple discriminate analyses (MDA). He defines the MDA approach as 'a statistical technique used to classify an observation into one of several a priori groupings dependent upon the observation's individual characteristics. It is used primarily to classify and / or make predictions in problems where the dependent variable appears in qualitative form'.

Thus, the MDA analysis uses a mix of fixed ratios and combines them with fixed coefficients. The result is a value which should have enough explainable power to describe company's current and future performance. According to Altman the linear MDA function follows the form:

$$Z = v_1x_1 + v_2x_2 + \dots + v_nx_n \quad (1)$$

where v_1, v_2, \dots, v_n = Discriminate coefficients
 x_1, x_2, \dots, x_n = Independent variable

Altman models the function with a relative small number of selected measurements. The underlying data set includes an overall sample size of 66 firms, splitted up into two sub groups consisting of 33 observations each. Whereas the first group includes firms following the characteristic non-bankrupt, the second group follows the characteristic bankrupt.

Nonbankrupt firms are chosen on a random base under the limitation that they were active in the business field production and that they have an asset size between \$1 and \$25 million. This justification is done to minimize the size effect. The bankruptcy sample covers a time horizon from 1946 to 1965. All analyzed bankrupt firms were active in the business field production as well as under the partition of chapter seven of the National Bankruptcy Act (US). His sample size is due to a lack of data for bankrupt companies with an assets size less than million and a rare bankruptcy possibility of larger firms. Altman collects the bankruptcy data from financial statements published one period before the bankruptcy takes place. Using this data, Altman uses t-statistics in order to find, which ratios are appropriate in order to forecast a bankruptcy. Based on that, he develops weights for the financial ratios to distinguish between defaulting and non-defaulting companies. Therefore, the Z-score function is described as:

$$Z = 0.012X_1 + 0.014X_2 + 0.033X_3 + 0.006X_4 + 0.999X_5 \quad (2)$$

- where
- X_1 = Working capital / Total assets; this ratio measures the net liquid assets relative to the firms total capitalization
 - X_2 = Retained earnings / Total assets; describes the cumulative profitability over time. According to Altman, the ratios advantage is that it captures the impact of a companies' age.
 - X_3 = Earnings before interest and taxes / Total assets; the ratio gives an impression of the productivity of the firms assets after tax and leverage factures.
 - X_4 = Market value of equity / Book value of total debt; the ratio describes the amount of value the companies assets can decline, before the liabilities exceed the assets.
 - X_5 = Sales / Total assets; Is the capital-turnover ratio, which describes the sales generating effect of firm's assets. According to Altman, it measures the managerial capability in dealing with competitive conditions.

For the underlying data set he finds that companies with a Z-score less than 1.81 the default occurs within one year. In contrast, firms with a Z-score exceeding 2.99 are solvent within the next year. Altman describes that the best cut-off value falls between 2.67 and 2.68 so that he defines the ideal Z-score value as 2.675. Applying the MDA function, Altman finds that it classifies 95% of all observations correctly. According to Richling et al. (2006), the model cannot be implemented into European solvency forecasts without changing the models weights. This is due to different accounting stands between the US and the EU.

2.2. Bounded Logit Model

In comparison to the linear Z-score model, the bounded Logit model uses non-linear techniques to compute the probability of default. Therefore, the models procedure is as follows; a firm can either go bankrupt or stay healthy, which can be described as for a bankrupt firm and for a non-bankrupt firm. The models probability that x is a defaulting company of the function i x can be described as:

$$P(Y_i = 1 | x_i) = P^*(\theta, x_i) = P_i^* \quad (3)$$

- where
- P = Probability
 - x_i = Regressor

The model aims to estimate. Therefore Cramer (2007) describes that the probability that a bankrupt firm is in a random sample follows the Bayes rule, which can be expressed as:

$$P_i = \frac{P_i^*}{P_i^* + \gamma(1 - P_i^*)} \quad (4)$$

where P = Probability
 γ = Fraction of non-bankrupt firms

According to Cramer, formula four can get maximized in order to find a correct decision. Therefore he describes, that if the fraction of non-defaulting firms is known, parameters of can be estimated from a given sample by using standard Maximum Likelihood methods. Using a sample size of 20,000 observations, Cramer finds the standard bonded model to forecast a firm's health best. Its main advantage against the standard Logit models is that its upper bound decreases the influence of outliers. The standard bounded Logit model follows the form:

$$w^*(\log[p/(1-p)]) = w^*(\beta_0 + \beta_1 y_1 + \beta_2 y_2 + \dots + \beta_i y_i) = w \frac{\exp(y_i^T \beta)}{1 + \exp(y_i^T \beta)} \quad (5)$$

where $1 - p$ = Probability that a firm defaults
 β_0 = Intercept term
 β = Coefficient associated with corresponding variable (whereas $i = 1, \dots, n$)
 w = Upper bound

Using binary dependent variables, the bounded Logit model estimates default probabilities relative to a cut-off value. Comparing the MDA approach to the Logit model, Tang (2006) describes the Logit models advantage over the MDA method is, that it does not assume multivariate normality. Whereas both models have in common that they use weighted ratios as input variables. Cramer defines the approach as an analysis that links the probability of a firm going bankrupt to its initial ratios. Like Altman, Cramer defines ratios as well as weights for his function, by analyzing which ratio in combination with which weight is most appropriated, to distinguish between defaulting and non-defaulting firms in the underlying data set. Therefore, the bounded Logit model transfers the input data into a nonlinear form whereas its upper bound is 1.1. According to Cramer the upper bound reduces the impact of outlines in the rating results. Using his data set, he estimates the upper bound as a best practice value such that it fits the data base best. Therefore, Cramer defines the bounded Logit model as:

$$P_i = 1.1 \frac{\exp(0.9 * Y_1 + 0.8 * Y_2 + 1.6 * Y_3 + 1.1 * Y_4)}{1 + \exp(0.9 * Y_1 + 0.8 * Y_2 + 1.6 * Y_3 + 1.1 * Y_4)} \quad (6)$$

where

Y_1	=	Own capital / Total assets; the ratio measures the firm's solvency. Therefore, the ratio describes the firm's ability to meet long term obligations.
Y_2	=	Gross returns / Total assets; the ratio measures a firm's rentability
Y_3	=	Working capital / Total assets; measures the amount of the short term obligations, which can be severed from assets.
Y_4	=	Net cash flow / Interest payments; shows the ability to make interest / principal payments.

Applying this Logit function, the cut-off value lies according to Cramer between 0.03 and 0.1. Whereas values above 0.1 explain, that the company will not default during the next year. A rating outcome under 0.03 describes a high probability that a bankruptcy occurs within the next year.

2.3. The KMV Model

More recent models differ substantially from the Z-score and the Logit approach. Wahrenburg et al. (2000) describe that two different approaches are dominating the academical as well as practical world nowadays. These two approaches are:

- Asset-value-models and
- Loss-rate-models

As this paper focuses on linear, Logit and as well as asset – value – models, the loss – rate approach will not be discussed and is just mentioned in sense of completeness.

Asset-value-models are based on option model theory. As Moody's asset-value-model has a major impact on the rating market, we test its reliability.

The KMV Model has been developed by KMV Corporation in 1988 and got sold to Moody's Corporation in 2002 (Moody's presents and promotes the model under www.moodyskmv.com). In this approach, the entrepreneur has an interest to liquidate the company if the market value of the firms' debt is higher than the market value of its equity. Following Sreedhar/Shumway (2004) the model computes this event by subtracting the face value of the firm's debt from an estimated market value of the firm and then divides this difference by an estimated value of the firm's volatility. This procedure results in a 'Z-score' value, which is expressed as the distance to default

(DD). In order to estimate whether the face value of a firm's debt is higher than the market value of its equity, the DD gets substituted into a density function.

To compensate the lack of unknown variables in the density function we use Sheedhar's naïve version of the KMV model. Sheedhar develops naïve probabilities to estimate the DD. Starting to measure a firm's market value, he assumes that the firm's market value equals the sum of its market value of debt and market value of equity. In order to estimate the market value of debt, the model sets the book value of debt equal to the market value of debt, which is defined as:

$$D = F \quad (7)$$

where D = Book value of debt
 F = Face value of debt

Furthermore, he assumes:

- that a company which is close to default has risky debt and
- that this risky debt correlates with risky equity.

Using this correlation the debt's volatility is assumed to be:

$$\sigma_D = 0.05 + 0.25 * \sigma_E \quad (8)$$

where σ_D = Variances of debt
 σ_E = Variances of equity

According to Sreedhar, the five percent value in equation eight represents the term structure of volatility. 0.25 times the equity volatility gets included to embrace the volatility associated with default risk. Combining equation seven and eight, Sreedhar describes the firms overall volatility as:

$$\sigma_V = \frac{E}{E+D} \sigma_E + \frac{D}{E+D} \sigma_D = \frac{E}{E+F} \sigma_E + \frac{F}{E+F} (0.05 + 0.25 * \sigma_E) \quad (9)$$

where D = Book value of debt
 F = Face value of debt
 σ_D = Variances of debt
 σ_E = Variances of equity

After computing the overall volatility, the companies expected return on assets equals the firm's stock returns over the previous year, which gets expressed as;

$$Naive\mu = r_{t-1} \quad (10)$$

where μ = Expected return on firms assets
 r_{t-1} = Stock returns over the previous year

According to Sreedhar combining equations nine and ten, the DD gets computed as:

$$DD = \frac{LN\left[\frac{(E+F)}{F}\right] + (r_{t-1} - 0.5\sigma_v^2)T}{\sigma_v\sqrt{T}} \quad (11)$$

where T = Forecasting horizon (=1)
 E = Equity
 F = Face value of debt
 σ_v = Firms overall volatility

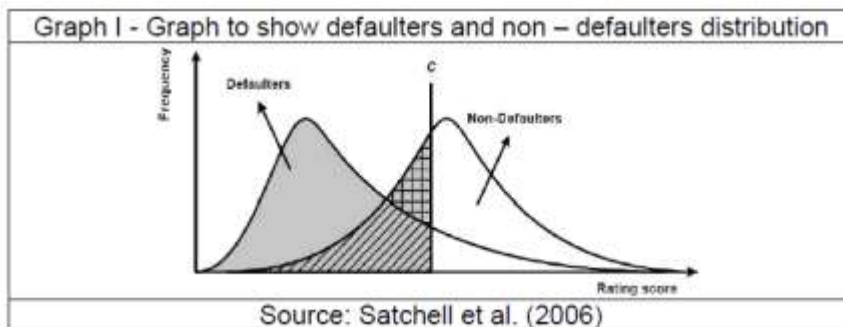
Furthermore, Sreedhar (2004) describes that the stock market data needs to be adjusted by the return of the firm in year t-1 minus the value-weighted S&P 500 / S&P t-1 (-).

As it is our aim to test models which are adaptable in the Basel Accord framework, the naïve version estimated DD's have to be as close as possible to the results estimated by the original KMV model. Sreedhar therefore mentions, that the naïve version keeps the structure of the KMV Model in terms of its DD as well as the expected default frequency without solving equations simultaneously. Furthermore, the model includes nearly the same quantity of information as the original KMV model, except of the underlying distribution and the market value of the firm's assets. According to Saudners et al. (2002), especially the underlying distribution highly influences the estimated DD. To bridge the unknown distribution function KMV applies, the naïve version assumes the underlying default function to be normally distributed. Furthermore, critical inputs used by the naïve version are the face value of equity and thus the equities volatility as the time horizon of the analyzed equity returns only covers one year. Sheedhar explicitly mentions that the model links the firm's equity value to a default event such that a declining equity value implies an increasing probability of default. Comparing this approach with the Z-score and the bounded Logit approach, the naïve KMV model reacts in the moment of declining share prices by scaling down the DD. The other two models require mainly balance sheet data, which gets published on a quarterly base. After computing the naïve KMV model and comparing the outcome with public available rating outcomes from KMV, Sreedhar concludes that the naïve KMV model has predictive power for default forecasts, whereas the models main strength is based on its functional form rather than from solving the two nonlinear equations.

3. Methodology

By their very own nature, rating models can be erroneous. Applying statistical tests in order to analyze ratings accuracy, Satchell et al. (2006) describe that any rating model needs to identify defaulting obligators from non-defaulting operators within a predefined time horizon better than a random model would do it. According to Beling et al. (2005) the models cut-off value therefore plays a crucial role in order to get an appropriate performance forecast. The cut-off value acts as a decision maker to classify an obligator. Altman's cut-off value for example is 2.675. Thus, if the Z-score model and its cut-off value are appropriate, it has to determine non-defaulting obligators and defaulting obligators with a higher likelihood as a random model would do it.

The situation, in which a rating system does not perfectly reflect reality, is described in graph I. The graph shows a distribution of defaulters as well as a distribution for non-defaulters. If the rating would perfectly distinguish between defaulting and non-defaulting firms in respect to the cut-off value, the two distributions would not overlap each other. Line C labels the models cut-off value.



If misevaluation takes place, two different types of errors can occur. Spuriously a company gets evaluated as a default candidate whereas the firm is healthy. Alternatively the firm could get misleadingly evaluated as a healthy one. Taking errors and correct decisions into account, table I presents all possible rating outcomes.

Table I - Table to show possible cut-off value scenarios	
Rating outcome is above the cut-off value and the company defaults	Rating outcome is above the cut-off value and the company does not default
- <i>Wrong decision</i> -	- <i>correct decision</i> -
- <i>α error</i> -	
Rating outcome is below cut-off value and the company defaults	Rating outcome is below the cut-off value and the company does not default
- <i>correct decision</i> -	- <i>Wrong decision</i> -
	- <i>β error</i> -

Blöchlinger et al. (2005) describes, that an alpha error occurs if the model estimates a lower risk as it is given. In contrast, the beta error describes, that the model estimates a company at a higher risk level as it is given in reality. In order to test ratings whether they reflect reality in a way which is sufficient to determine company's economical robustness, the Basel Committee on Banking Supervision (BCBS) (2000) published several approaches achieving these requirements. According to the BCBS, ROC and CAP curve approaches are an appropriate statistical measure to test rating accuracy. Satchell et al. (2006) refer to the BCBS (1999) by writing, that both methods are popular measurements to evaluate a rating models performance in practice. As an advantage of the ROC and CAP techniques against other performance measurements, Blöchlinger points their ability to visualize a systems performance. Therefore, the ROC and CAP graphs label the coordinate axis with hit and false rate.

3.1. The ROC curve

The ROC curve is defined by Blöchlinger et al. (2005) to be a 'two dimensional measure of classification performance and visualizes the information from the Kolmogorov – Smirnov statistics'. According to Engelmann et al. (2003) the ROC is computed by using the percentage of defaulters whose rating scores are equal or lower than the maximum score fraction of the overall sample size. Thus, the systems correctness is getting measured by using the total number of observations and the fraction of observations the system incorrectly assigns as non-defaulters. Starting with non-defaulters, their fraction is mathematically measured and expressed in terms of the hit-rate. According to Blöchlinger the hit-rate is described to be one minus the alpha error under the null hypothesis that high scores are translated into high default probabilities. Therefore, the hit rate gets expressed as:

$$HR(C) = \frac{H(C)}{N_D} \quad (12)$$

where $H(C)$ = Total number of defaulters, which were calculated correctly with the cut-off value C
 N_D = Total number of defaults in the sample

Thus, this measure describes the number of defaulting firms found correctly in the sample.

After estimating the number of defaulters found correctly in the sample, the amount of defaulters that were classified erroneously has to be identified. In order to do so, the false alarm rate has to get computed. According to Satchel et al. (2006), the false alarm rate is defined to be the number of non-defaulters that were classified

incorrectly as defaulters by using the cut-off value. Thus the false alarm rate measures the beta error. The rate is defined as:

$$FAR(C) = \frac{F(C)}{N_{ND}} \quad (13)$$

where $F(C)$ = Total number of non-defaults which were classified incorrectly as defaulters according to the cut-off value
 NN_D = Total number of non-defaulters the sample size

To illustrate that, we apply the HR and FAR methodology at the Altman (1968) paper. According to Altman (1968), the Z-score model has a targeting precision of 95%. Thus for Altman's data set, the following numbers of correct and incorrect observations are described:

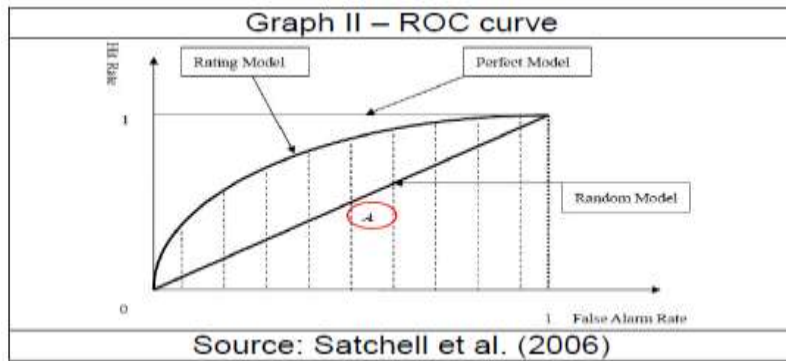
Table II Altman Model example		
Actual Group Membership	Predicted Group Membership	
	Bankrupt	Non-Bankrupt
Bankrupt	31	2
Non-Bankrupt	1	32

Therefore, the hit- and false alarm rates in respect to the cut-off value are:

$$HR(2.675) = 31 / 33 = 0.939$$

$$FAR(2.675) = 2 / 33 = 0.061$$

The actual ROC curve now is estimated by using all possible cut-off values given by the model and combining them with the rating outcomes to compute the HR and FAR rate. We apply cut-off values according SPSS. The program starts by using one plus the highest rating grade given by a model and goes down in 0.05 steps to one minus the lowest rating grade given by a model. Going back to the Altman example, the fraction of bankrupt to non-bankrupt firms differs, if another cut-off value is set. Having these values, the HR and FAR differ as they analyze the fraction of defaulters to non-defaulters in respect to the actual cutoff point. The results are then plotted in a graph. Graph II illustrates that.



As it can be seen from the graph II, the ROC curve's abscissa is labeled as false alarm rate and the ordinate labels the hit-rate.

According to Satchell, a rating models performance is better the steeper the ROC curve is at its left end and the closer the curves position is to the point (0.1). Thus, a models performance can be measured in terms of the area under the curve – the larger the area under the ROC the better the rating model. The area under the ROC (AUROC) is labeled as 'A' in graph II. According to Hutchinson (2005) the ROC approach follows two hypothetical Gaussian distributions.

Mathematically, the AUROC can be expressed as:

$$A = \int_0^1 HR(FAR) d(FAR) \quad (14)$$

where HR = Hit rate
 FAR = False alarm rate

As it is our aim and objective to find the rating system which offers the best performance in order to forecast bankruptcy, the decision rule is as follows: the system which produces the largest significant 'A' value is the one with the best performance.

3.2. The CAP curve

The CAP approach is alike the ROC approach. It is also used to measure the rating models performance. Instead of plotting the hit against the false rate like the ROC does it, the CAP uses the fraction of defaulters and plots it against the fraction of all obligators.

Satchell et al. (2006) define the CAP techniques as: 'for a given fraction x of the total number of debtors the CAP curve is constructed by calculating the percentage d(x) of the defaulters whose rating scores are equal to or lower than the maximum score of fraction x'. Graphically, we described the CAP curve in Graph III.

In order to apply the CAP, the rating outcomes have to be rearranged first. Therefore, the outcomes have to be ordered from the safest to the most risky obligator. A perfect rating model would allocate the lowest rating grade to the firm which is most likely to go bankrupt and the highest rating grade to the safest firm in the defined time horizon. In this case, the model would reflect reality perfectly and the CAP curve would go straight to the point (0.1) and stay at the line 1.1. In contrast, a random rating model is assumed to have no discriminate power. The random model is shown in graph III with the 45 degree line from point 0.0 to point 1.1. A real world rating scenario gives an output anywhere between a perfect and a random rating model.

The random rating line plays a crucial role for the evaluation of a rating model with the CAP technique. Like the ROC, the CAP uses the area under the curve as an assessment factor. According to Fernandes (2005), in comparison to the ROC, the CAP does not use the whole area under the rating model curve, but the area between the random rating model and the rating model curve as an estimator for the models performance. This area is described as the Accuracy Ratio (AR) and is defined to be:

$$AR = \frac{a_R}{a_P} \quad (15)$$

$$\begin{aligned} \text{where } a_R &= N_{ND}(A - 0.5) & \text{where ND equals the total} \\ a_P &= 0.5N_{ND} & \text{number of non-defaulters} \end{aligned}$$

Thus the decision rule for the CAP is; the larger the area between the random model curve and the rating model curve, the better the model describes reality. A perfect model therefore would be visualized in the graph as a horizontal line which crosses the point 1:1 in the coordination system.

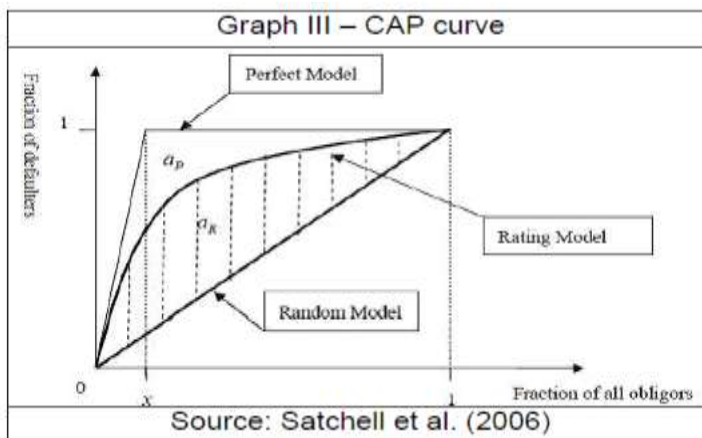
In order to represent a complete picture of the academical discussion linked to these two methodologies, Blockwitz et al. (2004) discusses problems of interpreting the ROC and CAP curves. Especially, the random model used in the CAP is critical in terms of describing the discriminate power of a model. Blockwitz focuses on the maximum value of one as a benchmark. Following their line of argumentation, this value would only occur, if all debtors are ranked correctly in relation to the random default event. This implies that after estimating a model results have to be ordered according to their value. Despite this critique, we use both models because the BCBS labels them as an appropriate tool to measure rating systems performances.

4. Data

Every default model we analyze forecasts the default probability for a time horizon of one year. Thus, if a company gets bankrupt in 2007, any rating model

should evaluate the company as a default candidate with data from 2006. In order to test the models, we use the most recent and largest corporate bankruptcies between 2006 and 2008 in the US. We built two data sets, one with annual and one with quarterly data, which provides us in total with 132 observations. According to the sample size of the annual and quarterly data sets, we imitate Altman's (1968) approach. That gives us a sample size of 66 observations for each set, splitted in 33 observations following the characteristics bankrupt and 33 observations following the characteristic non-bankrupt. Starting to collect bankruptcy data, we use the database bankruptcydata.com. As a matter of particular interest, the page offers names of the 20 largest US bankruptcies of each year. Hereby, the authors do not distinguish, whether the company got under chapter seven or 11 of the US bankruptcy code. Using these companies as a starting point, we get balance sheets, income statements, cash flow statements, as well as stock market prices from Google-finance, Yahoo-finance, Data stream as well as the Securities and Exchange Commission (SEC) database.

If it is not explicitly mentioned, the following paragraph does not distinguish between bankrupt and non-bankrupt companies. Especially the bounded Logit model and the KMV are models which are not restricted to a specific branch or business field. To test their broad applicability we collect data from firms, which are active in the following branches: Constructing companies (4), manufactures (8), energy production (2), telecommunication and information technology (5), retail industry (7), financial industry (5) as well as the airline industry (2). Furthermore, all companies have in common that they were / still are publicly traded. The accounting data provides all information to solve the bounded Logit model directly. In order to estimate the Altman model,



the equity market value is estimated by multiplying the amount of outstanding shares times the stock market price at the announcement day of the annual / quarterly report. Data which is used to solve the KMV model differs substantially from data we use to

solve the Z-score and bounded Logit model. According to the models description, it is necessary to solve the following variables: volatility of stock returns, the face value of company's debt and the standard deviation of return. Stock market data to estimate the value of equity - formula nine and 11 - is gained from Google finance and Data stream. The numbers of outstanding shares are published within the company's balance sheets, downloadable at SEC. According to Sreedhar et al. (2002), we substitute the face value of a debt with the total amount of debt plus current liabilities published in the balance sheet. According to the relative short time horizon of our data set, we do not adjust the data from outliers. Using this data in order to compute the default probabilities, panels A – D present descriptive statistics for the Altman and the Logit models. Panel I – J include descriptive statistics for the KMV model. The structure is as follows; first we present the insolvent yearly data and the solvent yearly data. This is followed by the insolvent quarterly and the insolvent yearly data description.

Facing Panel A and B we find that insolvent firm's have less debt than solvent firms. The higher standard deviation in the insolvent set indicates that debt is more dispersed for insolvent firms than for solvent firms. The EBIT draws a clear picture between solvent and insolvent firms. Whereas insolvent firms have on average a negative EBIT, solvent firms show a positive one. Comparing the EBIT maximum values, the insolvent data set shows still a positive value, but compared to the solvent sample it is more than 10.5 times lower. The market value of equity equals outstanding shares times the stock market price – both collected at the announcement day of the annual reports. Whereas the insolvent data set clearly indicates that the market evaluates defaulting firms low, the values of the solvent firms are highly dispersed among the sample. Thus, it seems reasonable, that the insolvent maximum value is by far lower than the minimum value of the solvent data set.

As the solvent firms EBIT value is higher than the one of insolvent firms, it seems realistic that retained earnings of insolvent firms are lower than retained earnings of solvent firms. On the other hand we observe that the insolvent values are not that dispersed among the sample as they are in the solvent data set.

Interesting to observe is, that insolvent firms do have much better sales values than solvent firms. Even though the insolvent data set shows a much higher standard deviation, both, the maximum as well as minimum values are higher than the comparable values in the solvent data set. Facing the equity book values, we also find that insolvent firms have higher equity book values than solvent firms have. Whereas both standard deviation values are very high, the one of the insolvent set is much higher than the solvent one.

In terms of gross returns, insolvent statistics are showing values far below the solvent values, whereas the solvent data set shows a higher standard deviation. As it can be assumed from the different debt levels, insolvent firms have less interest payments than solvent firms. Remarkable for both samples is the negative net cash flow whereas the solvent samples standard deviation is higher.

Both, total assets as well as working capital are used to estimate the Altman as well as the Logit model. Insolvent firms have on average less total assets and show a lower standard deviation than the solvent firms. Facing working capital we find that insolvent firms have on average more working capital than solvent firms. Furthermore, they show a higher standard deviation.

Turning the view to the quarterly data set we get comparable results as we get them for the yearly data set. Therefore, we focus on differences between the yearly and the quarterly data set. Descriptive statistics of the quarterly set are presented in Panels C – D. While the yearly set shows on average a positive EBIT for insolvent as well as for solvent firms, the quarterly set shows a highly negative EBIT mean for insolvent firms. Furthermore, interest payments of insolvent firms doubled for the annual data in comparison to the quarterly data set. Moreover, firms in the solvent sample generate a positive net cash flow whereas insolvent firms generate a much lower net cash flow than they do it in the annual sample. On the other hand, solvent firms show a negative mean of working capital what differs to insolvent firms, which are generating a positive working capital on average.

While presenting these results, it is worth to mention that in ten times, the last annual reports were closer to the default event than the last quarterly report were. Furthermore, we find that banks have a massive impact on the statistics presented. Descriptive statistics of the KMV model are presented in Panel G and J. We find that insolvent firms show lower equity volatility than solvent firms. Furthermore, insolvent firms have a lower debt volatility than solvent firms. In all samples, the debt volatility on average has a value between 0.59 and 8.15. Also interpreting mean values, we find that the overall volatility differs among the samples with value between 8.15 and 1.6., so that solvent firms have compared to insolvent firms higher values. In terms of returns, we find that insolvent firms have on average highly negative values whereas we find for solvent positive returns on average.

Panel A: Insolvent yearly

	Debt*	EBIT*	Equity market value*	Retained earnings*	Sales*	Equity book value**	Gross returns**	Interest payments**	Net cash flow**	Total assets***	Working capital***
Mean	936284	1394545	6.651265	3943.569	2587209	31273968	1211840	-1191.462	-1242.799	2980038	5329264
Median	271.135	4.045	1.092	-5.135	2017.08	49.07855	156.1205	-75.225	-0.38	2677.495	22.315
Maximum	22430000	49353259	28	981112	86376259	8.88E+08	27344000	76000	15798	96091001	1.77E+08
Minimum	0.3	-1252648	0.004	-628120	1.58	0.14576	0.112574	-50754	-40989	4.77	-778071
Std. Dev.	4009452	5479121	8.746704	206033.1	14805642	1.52E+08	4857544	17180.83	10060.43	16465006	30361899
Probability	0	0	0.037539	0	0	0	0	0	0	0	0

Panel B: Solvent yearly

	Debt*	EBIT*	Equity market value*	Retained earnings*	Sales*	Equity book value**	Gross returns**	Interest payments**	Net cash flow**	Total assets***	Working capital***
Mean	21411547	3338996	38055250	5113306	2293525	38055.25	2654038	-391644.5	-16312.48	34276528	1265571
Median	9940300	1907560	12107159	4887000	1090000	12107.16	980234	-195957	55000	14820700	611000
Maximum	1.55E+08	20101000	3.71E+08	35666000	12599000	370584.6	35706000	0.01	1845000	2.71E+08	29531900
Minimum	1039100	-893458	1354655	-45907000	-1001437	1354.655	1095.278	-2064000	-2899400	1839713	-14929000
Std. Dev.	30730949	4348907	76449457	13233097	3259386	76449.46	6233820	539458	732339.9	52103978	6678433
Probability	0	0	0	0	0	0	0	0	0	0	0

* = Used for Altman model

** = Used for Logit model

*** = Used for Altman and Logit model

Panel C: Insolvent quarterly

	Debt*	EBIT*	Equity market value*	Retained earnings*	Sales*	Equity book value**	Gross returns**	Interest payments**	Net cash flow**	Total assets***	Working capital***
Mean	14060893	-1993746	4.547447	53412.46	610478.4	1.45E+08	849022.9	-1154406	-1145218	15034940	3593404
Median	4184.822	-656.045	1.425	-14.625	139.03	115.0292	96.51	-2130	0.01	10681	-7.38
Maximum	4.52E+08	190370	44.78	1234471	20823020	3.32E+09	24384000	10	215063.4	4.52E+08	1.31E+08
Minimum	2.14	-51547843	0.004	-636220	-403989	0.1458	-28.56	-25665284	-36881683	157.19	-5678350
Std. Dev.	77488646	8904987	8.493276	298877.2	3572852	5.80E+08	4202548	4743270	6668072	77431968	22564634
Probability	0	0	0	0	0	0	0	0	0	0	0

Panel D: Solvent quarterly

	Debt*	EBIT*	Equity market value*	Retained earnings*	Sales*	Equity book value**	Gross returns**	Interest payments**	Net cash flow**	Total assets***	Working capital***
Mean	22084040	973437.2	25050263	5844523	568635.2	25050.26	1182804	-135275	349110.2	2.48E+08	-1053248
Median	8271400	527000	10910640	4815585	201500	10910.64	682000	-71000	439970	14009000	602393
Maximum	1.60E+08	6820000	2.65E+08	45647000	4707000	265000.6	12824000	-0.01	6241000	7.28E+09	15717000
Minimum	1531546	-1969091	31968.83	-43084000	-1251647	31.96883	-3161000	-868000	-9544189	2416273	-30470164
Std. Dev.	32296602	1715578	47034907	14810216	1122836	47034.91	2452650	193701.1	2319277	1.25E+09	7521579
Probability	0	0	0	0.00001	0	0	0	0	0	0	0

* = Used for Altman model

** = Used for Logit model

*** = Used for Altman and Logit model

Panel E: Altman Z-score Model - Descriptive Statistics

	Z-score Insolvent Yearly						Z-score Solvent Yearly					
	X1	X2	X3	X4	X5	Z-Score	X1	X2	X3	X4	X5	Z-Score
Mean	-0.083	-0.006	0.097	0.104	5.072	5.183	0.001	0.005	0.008	0.014	0.156	0.184
Median	0.000	0.000	0.001	0.002	1.113	1.115	0.000	0.004	0.004	0.008	0.074	0.090
Maximum	0.512	0.009	2.450	0.849	74.207	76.086	0.017	0.039	0.039	0.075	0.673	0.765
Minimum	-1.957	-0.071	-0.633	0.000	0.003	-0.301	-0.009	-0.067	0.000	0.001	-0.004	0.002
Std. Dev.	0.399	0.016	0.527	0.244	15.835	16.049	0.004	0.019	0.010	0.016	0.189	0.207
Jarque-Bera	-3.476	-2.929	3.534	2.515	3.771	3.795	27.477	329.201	23.385	58.848	15.307	15.312
Probability	16.460	11.291	15.351	7.801	15.566	15.817	0.000	0.000	0.000	0.000	0.000	0.000

X1 = Working capital / Total assets

X2 = Retained earnings / Total assets

X3 = EBIT / Total assets

X4 = Market value of equity / Book value of debt

X5 = Sales / Total assets⁴

Panel F: Logit model - Descriptive Statistics

	Logit Yearly Insolvent					Logit Yearly Solvent				
	Y1	Y2	Y3	Y4	Logit	Y1	Y2	Y3	Y4	Logit
Mean	0.133	0.108	-0.018	-0.096	0.548	0.012	0.093	0.418	2.211	0.644
Median	0.029	0.060	0.000	-0.069	0.542	0.002	0.059	0.047	0.372	0.730
Maximum	0.713	0.542	0.752	0.994	0.944	0.096	0.455	10.400	30.912	1.100
Minimum	0.000	0.001	-1.055	-1.665	0.000	0.000	0.002	-1.330	-6.275	0.005
Std. Dev.	0.190	0.128	0.334	0.485	0.195	0.026	0.096	1.869	7.529	0.361
Jarque-Bera	29.071	54.062	12.396	10.337	0.870	69.901	48.682	869.929	88.513	2.582
Probability	0.000	0.000	0.002	0.006	0.647	0.000	0.000	0.000	0.000	0.275

Y1 = Equity book value / Total assets

Y2 = Gross returns / Total assets

Y3 = Working capital / Total assets

Y4 = Net cash flow / Interest payments^a

Panel G: KMV Model - Descriptive Statistics

	KMV Insolvent Yearly					KMV Solvent Yearly				
	Volatility Equity	Volatility Debt	Vol Over*	Return	DD	Volatility Equity	Volatility Debt	Vol Over*	Return	DD
Mean	5.695	1.474	3.072	-109.651	-51.156	7.492	1.923	5.116	0.216	-2.236
Median	4.286	1.122	1.652	-1.300	-3.663	5.653	1.463	3.590	0.178	-1.525
Maximum	44.553	11.188	11.492	1,000,000	142.845	22.663	5.716	19.915	1.532	0.506
Minimum	0.000	0.050	0.011	-1914.882	-500.128	1.764	0.491	0.811	-0.269	-9.789
Std. Dev.	8.199	2.050	3.214	340.668	125.976	4.941	1.235	4.000	0.336	2.135
Jarque-Bera	330.442	330.442	5.131	848.092	56.674	9.391	9.391	42.566	56.485	28.069
Probability	0.000	0.000	0.077	0.000	0.000	0.009	0.009	0.000	0.000	0.000

* = Volatility Overall

Panel H: Altman Z-score model - Descriptive Statistics

	Z Score Insolvent Quarterly						Z Score Solvent Quarterly					
	X1	X2	X3	X4	X5	Z-Score	X1	X2	X3	X4	X5	Z-Score
Mean	-0.274	-0.042	-0.020	0.165	0.174	2.246	0.001	0.002	0.001	0.013	0.012	0.029
Median	0.000	0.000	-0.004	0.002	0.006	0.209	0.000	0.005	0.001	0.008	0.017	0.030
Maximum	0.367	0.010	0.011	1.828	1.614	74.052	0.008	0.015	0.003	0.074	0.070	0.118
Minimum	-9.391	-1.276	-0.293	0.000	-0.644	-9.817	-0.006	-0.027	-0.007	0.000	-0.137	-0.130
Std. Dev.	1.638	0.222	0.055	0.391	0.394	13.017	0.003	0.009	0.002	0.016	0.037	0.045
Jarque-Bera	-5.466	-5.468	-4.061	3.007	1.550	5.298	1.093	87.719	274.349	58.038	158.193	35.623
Probability	30.937	30.944	19.891	11.874	7.031	29.863	0.579	0.000	0.000	0.000	0.000	0.000

X1 = Working capital / Total assets

X2 = Retained earnings / Total assets

X3 = EBIT / Total assets

X4 = Market value of equity / Book value of debt

X5 = Sales / Total assets^a

Panel I: Logit model - Descriptive Statistics

	Logit quarterly insolvent					Logit quarterly solvent				
	Y1	Y2	Y3	Y4	Logit	Y1	Y2	Y3	Y4	Logit
Mean	0.122	0.257	0.838	-0.008	0.571	0.010	0.054	-972.078	4.671	0.843
Median	0.055	0.023	-0.011	0.000	0.610	0.001	0.056	0.059	5.067	1.088
Maximum	0.619	5.050	30.615	0.632	1.100	0.311	0.205	1.025	23.695	1.100
Minimum	0.000	-0.098	-1.338	-0.971	0.000	0.000	-0.130	-32080.730	-6.034	0.000
Std. Dev.	0.157	0.886	5.359	0.338	0.257	0.054	0.060	5584.551	5.802	0.412
Jarque-Bera	18.690	979.418	1219.368	3.507	1.155	1244.223	6.409	1245.579	8.060	9.197
Probability	0.000	0.000	0.000	0.173	0.561	0.000	0.041	0.000	0.018	0.010

Y1 = Equity book value / Total assets

Y2 = Gross returns / Total assets

Y3 = Working capital / Total assets

Y4 = Net cash flow / Interest payments^a

Panel J: KMV Model - Descriptive Statistics										
	KMV Insolvent Quarterly					KMV Solvent Quarterly				
	Volatility Equity	Volatility Debt	Vol Over*	Return	DD	Volatility Equity	Volatility Debt	Vol Over*	Return	DD
Mean	2.186	0.557	1.601	-0.430	33.368	10.831	2.758	8.154	0.088	1.592
Median	0.952	0.288	0.351	-0.378	-0.663	9.165	2.341	3.658	-0.011	0.691
Maximum	11.476	2.919	11.365	0.238	1091.336	35.233	8.858	42.398	1.180	16.670
Minimum	0.000	0.050	0.007	-0.989	-12.370	2.166	0.592	0.209	-0.640	-4.312
Std. Dev.	3.084	0.771	2.871	0.387	193.122	7.064	1.766	9.927	0.433	4.414
Jarque-Bera	27.600	27.600	76.253	2.396	1125.950	20.778	20.778	32.354	6.526	32.230
Probability	0.000	0.000	0.000	0.302	0.000	0.000	0.000	0.000	0.038	0.000

* = Volatility Overall

5. Results

As it can be assumed while reading the models descriptive statistics, the three different models show an inhomogeneous performance. Whereas the performance varies among the different models it also differs among the data sets. While presenting the results we keep the structure used in the parts above, so that we first present the Altman model, which is followed by the Logit and the KMV model. After presenting descriptive statistics of the models results, we do not distinguish anymore between the solvent and insolvent data set but use the full data set to analyze the models performance. Descriptive statistics are followed by analyzes of alpha and beta errors. The evaluation is concluded by presenting the ROC and CAP results.

As it can be seen from Panel E, the yearly Z-score values for insolvent companies differ quite a lot among the sample. Whereas the mean is about 5.18, the minimum value is around -0.13 and the maximum value is about 76.08. Thus, according to the cut-off value of 2.675, we can assume that the model forecasts insolvent firms mainly incorrect. Furthermore, it is conspicuous that the standard deviation has, compared to the other models, the highest value with 16.049. The solvent data Z-scores differ sustainably from the insolvent Z-scores. Here, the mean is around 28.17 times lower as what we observe in the insolvent data set. Furthermore, both the maximum as well as the minimum values are below zero. Linking the values to the cut-off value, we find that the model cannot correctly forecast this sample. Coming to the yearly Logit model in Panel F, the statistics are painting another picture. Reminder; the model has an upper bound, means that the maximum values cannot extent 1.1. The insolvents set maximum value is close to the upper bound but does not reach it, as it has a value of 0.944. With a minimum value of zero, the model reaches a standard deviation of 0.195, which is compared to the other models the lowest value in the sample. Thus, we can confirm Cramer's (2007) observation, that the bounded Logit model reduced the occurrence of outliers. According to the cut-off value of 0.1, we find that the model evaluates defaulting firms mostly incorrect. The counter sample produces higher maximum as well as minimum values as the insolvent data produces. The maximum value reaches the models upper bound. Furthermore, the mean equals

0.73. According to the cut-off value we can assume that the model forecasts solvent firms mainly correctly.

Coming to the KMV models distance to default value in Panel G, we find highly dispersed values for the insolvent data set. As the standard deviation exceeds 125, the model generated the highest value in the sample. With a mean of -51.1 a maximum value of 142.84 and a minimum value of -500,128 we assume that the model forecasts the dominating amount of defaulting firms correctly. This picture changes by facing the solvent data. Here, the model generates values close to zero or even negative. With a maximum value of 0.5 and a minimum value of -9.7 the model generates a standard deviation about 2.14. Combining that with a negative mean value, we can assume that the model forecasts solvent firms mainly incorrect. Furthermore, for both samples, the model generates a distribution, which is not normally distributed.

Compared to the yearly data set, the quarterly data shows differences. The Altman model, presented in Panel H, estimates a maximum value of 74.05, a minimum value of -0.81, with a standard deviation of 13.017. Having a mean value of 2.246 and the cut-off value of 2.675 we get low evidence to presuppose in which direction the model could forecast firms. However, the mean gives evidence to assume that the model could forecast defaulting firms incorrectly. Estimations with solvent data generating values close to one. With a very low standard deviation of 0.045 and a mean of 0.029 we can assume the model to forecast defaulting firms mainly incorrect. The Logit's insolvent estimations in Panel I show a maximum value equal to the models upper bound and a minimum value of zero. As the standard deviation is around 0.25 and the mean is 0.57 we assume the model having problems to find defaulting firms. For the quarterly data, the model generates the same maximum and minimum value as it generates them with the yearly data, but reaches a higher standard deviation of 0.41, as well as a higher mean of 0.83. According to the cut-off value we can assume that the model mainly forecasts solvent firms correct.

Coming to the KMV quarterly results in Panel J, we find that that the model reaches a very high maximum value in combination with a relative small minimum value. As the standard deviation is also very high we can assume that the model generates outlines which are influencing the statistics. The solvent data differs. Here, the values are relatively small as the maximum equals 16.67 and the minimum equals -4.31, the model generates a standard deviation of 4.4. Having a mean of 1.59 we can assume that the model finds solvent firms better than it finds insolvent firms. Furthermore, the model does not generate a normal distribution, which is in the version we are using, a fundamental underlying assumption.

5.1. Alpha / Beta errors

Remembering that the alpha error describes that even though the rating outcome forecasts no bankruptcy, the firm defaults. In comparison, a beta error

explains that a rating assumes a firm to go bankrupt which is not the case in reality. Table III therefore summarizes our rating outcomes categorized into correct decisions and alpha and beta errors.

Table III: Alpha / Beta errors			
Yearly data N=66		Quarterly data N=66	
Altman model cut-off-value = 2.675			
Alpha error=6	Correct decision*=0	Alpha error=1	Correct decision*=0
Correct decision**=27	Beta error=33	Correct decision**=32	Beta error =33
Logit Model cut-off-value = 0.1			
Alpha error=32	Correct decision*=30	Alpha error=31	Correct decision*=28
Correct decision**=1	Beta error=3	Correct decision**=2	Beta error=5
KMV Model			
Alpha error=6	Correct decision*=1	Alpha error=12	Correct decision*=22
Correct decision**=27	Beta error=32	Correct decision**=21	Beta error=11
* = Value is above the cutoff value and the company does not default			
** = Value is below cutoff value and the company defaults			

As one probably presumes by reading the descriptive statistics, the Altman models performance is, in terms of the beta error and its correct decisions low. We find the alpha error six and the beta error 33 times. The model assumes every company in the solvent data set to be bankrupt. In total, the model does 27 correct decisions. Whereas its strength is to correctly find defaulting firms. Thus, out of 66 observations, the Altman model categorizes 40.9% of all firms correctly.

Interestingly, the model shows another performance for the quarterly data set. The alpha error decreases to one observation and the beta error stays constant at 33 observations. In total, 32 correct decisions are done. Out of the solvent data set, the model assumes no company to be solvent next year. Furthermore, it does 33 correct decisions by finding bankrupt companies. Thus, the Altman model evaluates 48.48% for the firms in the quarterly data set correctly. Noticeable for both data sets is that the model has a very high beta error and thus, it evaluates every solvent company incorrectly.

Coming to the yearly Logit model, it is obvious that it has problems identifying bankrupt candidates. As table III shows, the model produces 32 alpha errors. In comparison, the beta error is very low with a total of three observations. In sum, it does 31 right decisions, whereas it only has two correct observations in terms of an actual bankruptcy, so that it evaluates 28 firms correctly. Thus, the model finds in 46.96% of all observation the correct decision. The results changes slightly by analyzing the quarterly data set. Here, we observe an alpha error of 31 and a beta error of five observations. In 30 times the model finds the right forecast. Out of the 30 correct decisions it only forecasts two bankrupt firms properly. Thus, the model finds 45.45% of the firms in the data set correctly.

In comparison to the Altman model, which mainly forecasts default candidates correctly, the Logit model has a very high alpha error. Therefore, we summarize that the Logit model forecast defaulting firms mostly wrong, but forecasts non-bankrupt firms with a higher probability than the Altman model does it.

Below the Altman and the Logit model, table III presents alpha and beta errors done by the KMV model. Reminder, compared to the Altman and the Logit model, the KMV model is the only one, which is based on option pricing theory. Facing the yearly data set, the KMV model does the alpha error six and the beta error 32 times. In total, it finds 28 times the right decision, whereas it evaluates only one non-defaulting company correctly. Thus it does in 42.42 times the correct decision. The picture changes by coming to the quarterly KMV outcomes. Whereas we find 12 alpha errors, the beta error occurs 11 times. The total amount of correct decisions is 43. Thus, the KMV model finds in 65.15% of all observations the correct decision. Interesting to observe is that the models performances differs quite a lot. Whereas it shows a hit rate of 42.42% for the yearly model, the performance for the quarterly data set is about 29.24% better.

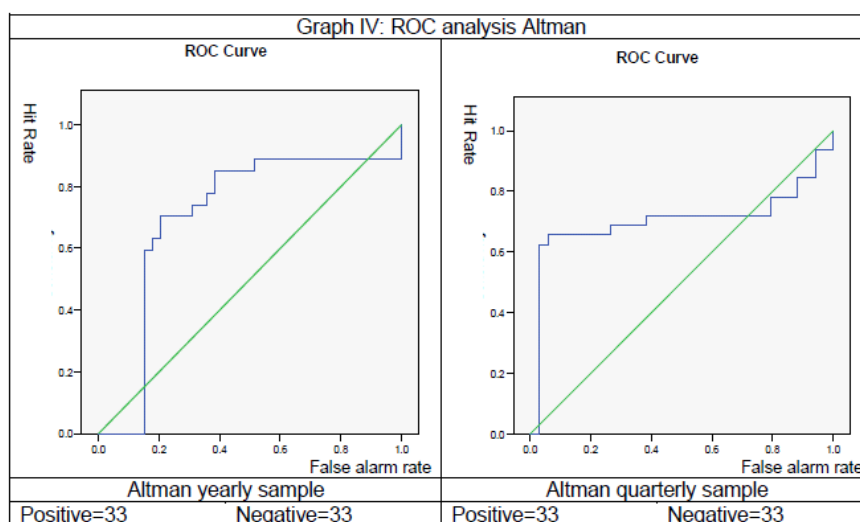
Even if these results do not give an interpretable outcome in terms of the Basel Accord framework, they indicate that all three rating models could have too high false rates. To do further analysis we now present the ROC and CAP estimations.

5.2. ROC results

While discussing the ROC curve and its AUROC, the decision rule is as follows. The larger the area under the curve the better is the ratings performance. Therefore, a value of one indicates, that the model has the highest possible explanatory power. A value of zero indicates that the model has no explanatory power. In a real world scenario, the models performance should be anywhere between zero and one. Reminder; by estimating the ROC, the program changes the cut-off values to get different hit- and false alarm rates. Therefore, it starts with one plus the highest rating grade as a cut-off value and goes down in 0.05 steps till one minus the lowest rating grade. Exemplary, we estimate the hit- false alarm rate with the models original cut-off values in table IV.

Table IV: Hit- False alarm rates			
Models / Data sets	Hit rate	False alarm rate	Σ
Altman yearly	0 / 33 = 0.000	33 / 33 = 1	1
Altman quarterly	0 / 33 = 0.000	33 / 33 = 1	1
Logit yearly	30 / 33 = 0.909	3 / 33 = 0.091	1
Logit quarterly	28 / 33 = 0.848	5 / 33 = 0.152	1
KMV yearly	1 / 33 = 0.031	32 / 33 = 0.969	1
KMV quarterly	22 / 33 = 0.666	11 / 33 = 0.334	1

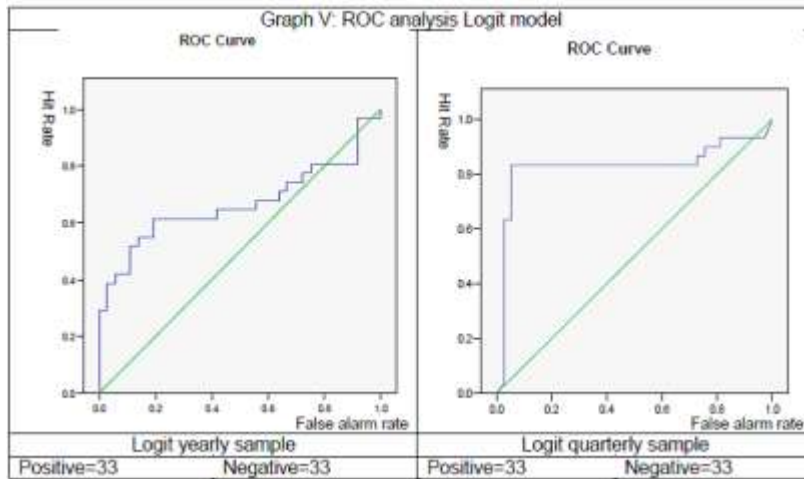
In the ROC graphs presented below, the 45 degree line labels a random models performance, the curve labels the ROC estimations. The significance values are based on a 95% confidence interval.



Graph IV shows the Altman models ROC estimations for both data sets. Both curves are created with 33 defaulting and 33 non-defaulting companies. The estimates are presented in table V. Altman's yearly AUROC equals 0.704. Furthermore, it generates a significance value of 0.05. Thus, the Altman model has an explanatory power of in terms of the significance value.

Table V.I Test Result Altman Yearly					Table V.II Test Result Altman Quarterly				
Area	Std. Error	Asymptotic Sig.(b)	Asymptotic 95% Conf. Interval		Area	Std. Error	Asymptotic Sig.(b)	Asymptotic 95% Conf. Interval	
			Lower Bound	Upper Bound				Lower Bound	Upper Bound
.704	.072	.005	.563	.844	.704	.074	.004	.560	.848
b. Null hypothesis: true area = 0.5					b. Null hypothesis: true area = 0.5				

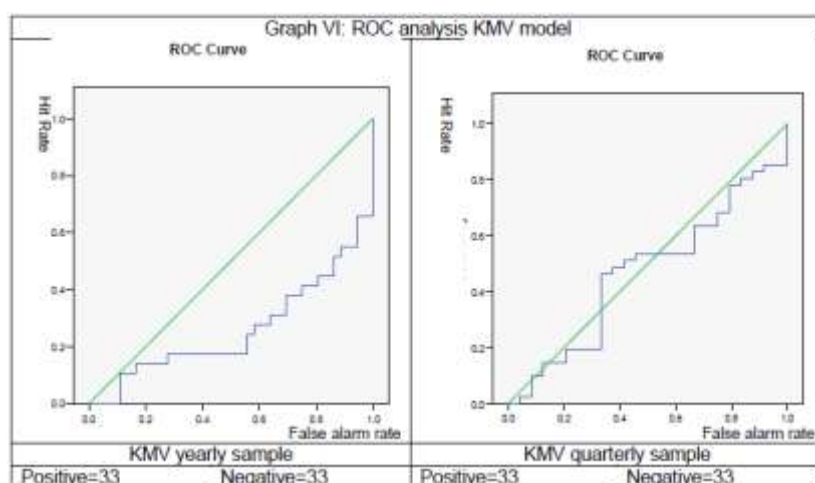
The models explanatory power slightly improves by analyzing the quarterly data set. Here, the significance value equals 0.04 whereas the AUROC stays constant at 0.704. Furthermore, the upper bounds level stays dominate at a nearly constant level. Summarizing the Altman model for both data sets by applying ROC techniques, we find that the models power for both samples is close to each other, whereas both models fulfill the requirements of the ROC to be better than a random model.



Next, we test the Logit model. As it can be seen on the left side of graph V - the yearly sample - a major part of the ROC curve, except of the curves upper right corner, is above the random models line. Therefore, table VI.I shows, that the AUROCs value is 0.667. As the significance value is above 0.05 the model has less explanatory power. Furthermore, the curve has a dominating upper bound.

Table VI.I Test Results Logit Yearly					Table VI.II Test Logit quarterly				
Area	Std. Error	Asymptotic Sig.(b)	Asymptotic 95% Conf. Interval		Area	Std. Error(a)	Asymptotic Sig.(b)	Asymptotic 95% Conf. Interval	
			Lower Bound	Upper Bound				Lower Bound	Upper Bound
.667	.072	.019	.526	.808	.830	.062	.000	.709	.951
b Null hypothesis: true area = 0.5					b Null hypothesis: true area = 0.5				

Looking at the right part of graph V we see that the ROC curve is highly above the random line. Only at the upper right part the curve is under the random models line. Table VI.II. presents the test results. As it can be seen, the AUROC equals 0.83 and the model produces a significance value of zero which proofs, that the model performs better than the random model does it. In comparison to the Altman and Logit model, the KMV model performs badly. It's clearly visible from Graph V.I that the yearly estimations performing worse than a random model would do it. Therefore, we estimate an AUROC value of 0.289. The models significance value is 0.01 and thus it describes that the model has explanatory power.



The quarterly ROC curve is close to the random model line. Therefore, the model produces an AUROC value of 0.467. Table VII.II shows the test estimations. We find a very high significant value of 0.664. Based on that, the model is described to have no explanatory power.

Table VII.I Test Result: KMV Yearly					Table VII.II Test Result KMV Quarterly				
Area	Std. Error	Asymptotic Sig.(b)	Asymptotic 95% Confi. Interval		Area	Std. Error	Asymptotic Sig.(b)	Asymptotic 95% Confi. Interval	
			Lower Bound	Upper Bound				Lower Bound	Upper Bound
.259	.064	.001	.132	.385	.467	.074	.664	.322	.613
b. Null hypothesis: true area = 0.5					b. Null hypothesis: true area = 0.5				

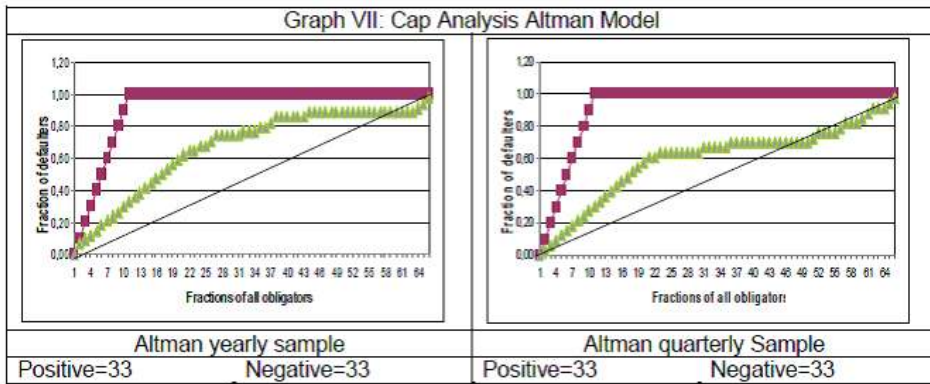
Summarizing the results based on the CAP estimations, we find that the Logit model performs best as it produces the highest significant AUROC value. The Logit model is followed by the Altman model, which generates two significant AUROC values. In comparison to the Logit model, both values are lower than the highest Logit AUROC. The KMV model performs worst; it generates the worst significant value and the smallest AUROC.

After presenting the ROC estimations we apply further analyzes by employing CAP techniques.

5.3. CAP Results

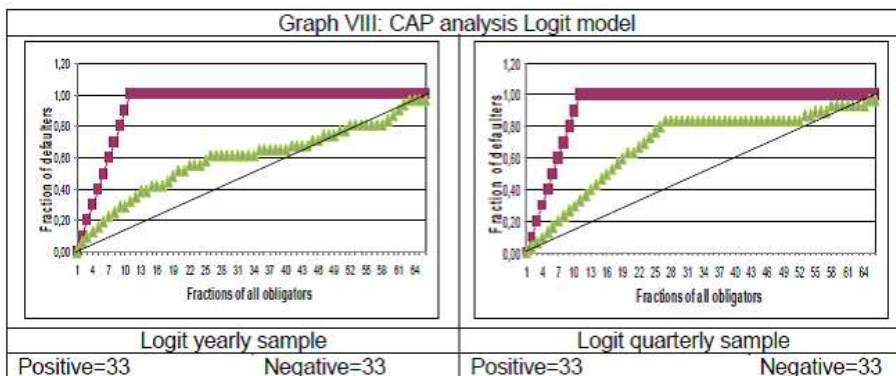
Our CAP estimations are based on a self made program. Even though the program is in line with comparable programs and computes the same test results, it does not provide a significance value as it was given for the ROC estimations. Therefore, the models evaluations are based on the integral between the random model and the CAP curve only. In all graphs, the line consisting of triangles (▲) labels the analyzed ratings. The line consisting of quadrates (■) labels the perfect rating

model and the 45 degree line (-) labels a random rating outcome. Here, the decision rule is as follows; the larger the area between the analyzed models and the random models line, the better the rating.



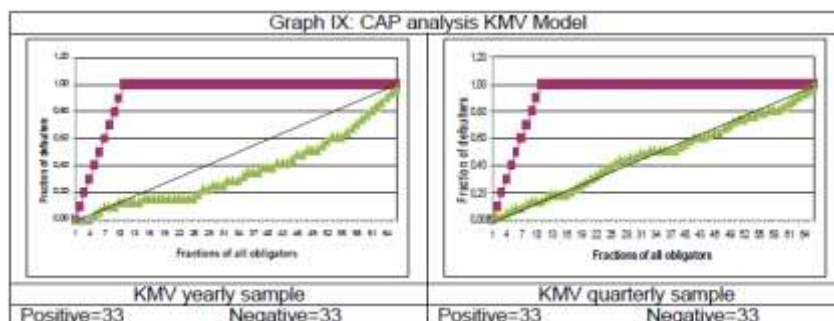
Graph VII displays the Altman models performances in terms of the CAP curve. The yearly sample produces a CAP curve which is, except of seven items in the upper right corner, mainly above the random models line. For the yearly data set, CAP estimations give a value of 0.32. Therefore, the model performs better than the random model. In its right part, graph VII shows the quarterly CAP curve of the Altman model. The graph shows that 20 items are placed under the random models line. The CAP estimator equals 0,172. Therefore we can summarize that the Altman model produces a curve, which is mostly above the random ratings line. Furthermore we find differences between the two data sets. The CAP estimator of the yearly sample is 1.86 times higher than the estimator for the quarterly sample.

Applying the same processes to the Logit model, we find a CAP value of 0.2992 for the yearly data set - Graph VIII shows the estimations. As it can be seen, 12 items are placed under the random models line. The quarterly data set generates a slightly better CAP value of 0.3059. Here four items are lying under the random models line.



The KMV models performance, measured in terms of the CAP, is bad. As it is visible from Graph IX, the CAP curve for the yearly data set is far below the random model and thus, it generates a negative CAP value of -0.52.

The quarterly data set generates a CAP curve, which is loosely to the random model line. Therefore, it produces a value of 0.0892, which is the lowest positive Cap value we measured.



As it is our interest to find the best performing model, table VII summarizes the ROC and CAP results. Table seven provides information to determine the most accurate model. Therefore, the model with the highest significant ROC and CAP values is placed first. The model, which generates the second highest significant ROC as well as the second highest CAP, is classified at place two and so forth. If two models have either a higher significant Cap value or a higher CAP value they both are classified with same rank.

Table VII: ROC and CAP ranking				
Rank	Model / data set	AUROC	Significance ROC	CAP
1	Logit (quarterly)	0.830	0.00	0.306
2	Altman (yearly)	0.704	0.05	0.320
3	Logit (yearly)	0.667	0.19	0.299
	Altman (quarterly)	0.704	0.04	0.172
4	KMV (quarterly)	0.467	0.664	0.089
5	KMV (yearly)	0.259	0.010	-0.520

As the models descriptive statistics and the analysis of alpha and beta errors already have shown, table VII concludes the picture that the models perform differently. Furthermore, we also observe differences between the yearly and the quarterly data set. Facing the ROC values, we find the Logit model to perform best as it produces the highest significant AUROC value in combination with the second highest CAP estimator. Thus, the Logit model is placed first. The Altman model also generates a significant AUROC value but compared to the Logit model the area under the curve is smaller, thus it takes place two. Rank three shows the Logit as well as the

Altman model. The Logit model has a higher CAP estimator whereas the Altman model has a significant AUROC value. The KMV model performs worst such that it reaches rank four and five. Whereas the yearly KMV estimations are generating the lowest significant AUROC value, it also generates a negative CAP value. The quarterly KMV estimations show an insignificant ROC value in combination with the CAP estimator, which is close to the random model.

6. Conclusion

We test the linear Altman model, the Logit model as well as the KMV rating model towards their performance. We analyze the Altman model as it is the first well known rating system, which is still used as a benchmark for current systems. Furthermore, we test a Logit model, as this type of model is used by J.P. Morgan, as well as the KMV model, which is used by Moody's. Both, Moody's KMV and the J.P. Morgan model have a major impact on the credit and rating market. In order to test these models, we use a US data set of 132 observations. The data is splitted into a yearly sample of 33 bankrupt and 33 non-bankrupt firms as well as a yearly sample including the same amount of observations. We measure the models performance in terms of the ROC and CAP techniques as they are applicable for the evaluation of rating systems within the framework of Basel Accords. The overall test results can be ranked as follows: the Logit model performs best, followed by the Altman model and the KMV model.

As the Logit model is one of the most applied models nowadays it seems reasonable that it outperforms the Altman model. Our results might surprise as the Altman model is the oldest model and originally developed in order to analyze US manufacturing companies only, but outperforms the KMV model. By interpreting the KMV results we can not negate an impact of the current crises on the stock market data. By interpreting the Logit models performance it is noticeable that it produces a very high alpha error. As the model only uses balance sheet data to estimate the Logit values we assume that even it is the best tested model its performance could be pushed by changing the models weights. Furthermore, it can be questioned whether the variable used in the model as sufficient in order to perform an accurate rating. The KMV model shows inconsistent patterns in its alpha and beta errors so that the models performance is far from being assumed to be an accurate rating method. In our framework we do not mimic the Moody's approach in total as it is unknown. We mimic an approach, which is described to estimate the distance to default as nearly adequate as the Moody's version did it in 2002. As we use later data Moody's model could be enhanced so that the current version of the model could estimate other results. Furthermore, we do not know, whether Moody's uses specific business related

indicators after computing the distance to default. In our approach we do not use any business field related indicators. Furthermore, stock market returns we used are mainly from 2007 and 2008. Even if we adjust the returns we can not exclude that the data is biased by the current crisis. Evidence, which supports the bad performance we find for the KMV model is, that several banks were rated with an AAA rating but were already affected by the crisis. The Hypo Real Estate bank for example had an Moody's AAA rating till it had to raise governmental credits in order not to get bankrupt. KMV downgraded the distance to default right after the bank was protected from getting bankrupt. Another explanation was given by Bjorn Stibbe, Senior Vice President of Rabobank for Leveraged Finance. He argued that ratings are negotiable. He mentions that especially large companies know about the inputs for a rating, such that they can influence data in order to get a better rating. Besides, firms pay fees to Moody's and J.P. Morgan in order to be rated. Therefore, firms could have an interest to influence the ratings towards a better result. A weakness of our research could be seen in the relative small sample size which could bias the results. Furthermore, we do not exclude financial institutions from the sample size as it is done in most of financial researched papers. Indeed, we observe that outliers are mainly caused by financial institutions. On the other hand rating models are getting developed in order to evaluate the firm independently from its business field.

Furthermore, the sample size we used could have an impact on the ROC curves. As we downsize the ROC graphs the curve gets more angulated as many rating outcomes are too close to each other that they cannot be drawn correctly. As that is only a mechanical problem, it has no effect on the estimations.

7. References

- Allen L. (2002): '*Credit Risk Modeling of Middle Markets*', presented at *Conference on Credit Risk Modelling and Decisioning*, Wharton FIC, university of Pennsylvania.
- Altman, E. I. (1968): '*Financial Ratios, Discriminate Analysis and the Prediction of Corporate Bankruptcy*', *Journal of Finance*, 589-609.
- Altman, E. (2002): '*Revisiting Credit Scoring Models in a Basel 2 Environment*', NYU Working Paper No. FIN-02-041.
- Arora, N. / Bohn, J. / Zhu, F. (2005): '*Reduced Form vs. Structural Models of Credit Risk: A Case Study of three Models*', Moody's I K•M•V•
- Barniv, R. / McDonald, J. B. (1999): '*Review of Categorical Models for Classification Issues in Accounting and Finance*', *Review of Quantitative Finance and Accounting*, Vol. 13, 39-62.
- Basel Committee on Banking Supervision (2000): '*Supervisory Risk Assessment and Early Warning Systems*'.
- Basel Committee on Banking Supervision (1999): '*Credit Risk Modeling: Current Practices and Applications*'.
- Beaver, W. H. (1966): '*Financial Ratios as Predictors of Failure*', *Journal of Accounting Research*, 4, 71-102.

- Beling, P. / Covaliu, Z. / Oliver, R.M. (2005): '*Optimal Scoring Cutoff Policies and Efficient Frontiers*', Journal of the Operational Research Society 56, 1016-1029.
- Blockwizt, S. / Hamerle, A. / Hohl, S. / Rauhmeier, R. / Rösch, D. (2004): '*Myth and reality of discriminatory power of rating systems*', Wilmott Magazine 2005, 2 -6.
- Blöchlinger, A. / Leippold, M. (2006): '*Economic benefit of powerful credit scoring*', Journal of Banking and Finance, Vol. 30 (3), 851-873.
- Cramer, J. S. (2007): '*Scoring bank loans, a case study that might go wrong*', Tinbergen Institute Amsterdam.
- Eisen, M. (2008): '*Haftung und Regulierung internationaler Rating-Agenturen*', Frankfurt.
- Engelmann, B. / Tasche, D. (2003): '*Testing rating accuracy*', www.risk.net.
- Fernandes, J. E. (2005): '*Corporate credit risk modeling, quantitative rating systems and probability of default estimation*', Lisbon University.
- Hutchinson, T. P. (2005): '*ROC analysis in credit scoring and credit judgment*', Database Marketing & Customer Strategy Management, Vol. 13, 3, 182–185.
- Merton, R. C. (1974): '*On the pricing of corporate debt: The risk structure of interest rates*', Journal of Finance 29, 449 - 470.
- Mitchell, J. / van Roy, P. (2007): '*Failure Protection Models: Performance, Disagreements and Internal Ratings*', National Bank of Belgium working paper no. 123.
- O.V. Frankfurter Allgemeine Zeitung (FAZ) (2008): <http://www.faz.net/s/RubDDBD ABB9457 A437 BAA85A49C26FB23A0/Doc~ E19B3E87FD51040B0B837F86D3A 7E20A6~ATpl~Ecommon~Scontent.html>. 01.10.2008
- O.V. Financial times Germany (FTD) (2007): http://www.ftd.de/boersen_maerkte/aktien/marktberichte/EZB%20Kontrolle%20Rating %20Agenturen/283076.html 23.11.2007
- O.V Manager Magazine (2008) : USA Gold (2008) <http://www.usagold.com/amk/usagoldmarketupdate082707.html> 01.09.2008 09.08.2008
- Reichling, P. / Dreher, D. / Beinert, C. (2006): '*Zur Verwendung des Altmannschen Z"-Scores als Benchmark für die Trennschärfe von Ratingfunktionen*', Otto-von-Guericke-Universität Marburg.
- Satchell, S. / Xia, W. (2006): '*Analytic Models of the ROC Curve: Applications to Credit Rating Model Validation*'.
- Saudners, A. / Allen, L. (2002): '*Credit Risk Measurement New Approaches to Value at Risk and Other Paradigms*', New York.
- Sreedhar, B. T. / Shumway, T. (2004): '*Forecasting Default with the KMVMerton Model*', Boston Meetings Paper.
- Tang, T. / Chi, LL. (2006): '*Bankruptcy prediction: Application of Logit Analysis in Export Credit Risk*', Australian Journal of Management, Vol. 31, No. 1.
- Wahrenburg, M. / Niethen, S. (2000): '*Vergleichende Analyse alternativer Kreditrisikomodelle*', Working paper series: finance and accounting.